Article Watch, April 2013

Clive Slaughter

GHSU-UGA Medical Partnership, Athens, GA

This column highlights recently published articles that are of interest to the readership of this publication. We encourage ABRF members to forward information on articles they feel are important and useful to Clive Slaughter, GRU-UGA Medical Partnership, 1425 Prince Ave., Athens, GA 30606; Phone: 706-713-2216; Fax: 706-713-2221; E-mail: cslaught@uga.edu; or to any member of the editorial board. Article summaries reflect the reviewer's opinions and not necessarily those of the association.

DNA SEQUENCING AND CHARACTERIZATION

Mathieson W, Thomas G A. Simultaneously extracting DNA, RNA, and protein using kits: is sample quantity or quality prejudiced? *Analytical Biochemistry* 433;2013: 10–18.

Kits for simultaneously extracting DNA, RNA, and protein from the same sample are available commercially. The present paper presents a performance comparison of two such kits, triplePrep from GE Healthcare and AllPrep from Qiagen. The kits are tested on good-, intermediate-, and poor-quality tissue samples, and their performance is compared with specialized kits for DNA extraction (Puregene from Qiagen), RNA extraction (RNeasy from Qiagen), and homogenization into buffer for protein extraction. Puregene DNA yields are 183% those of triplePrep and 506% those of AllPrep. RNeasy RNA yields are similar to those of AllPrep, but 412–588% those of triplePrep. The yield of protein by extraction into buffer is 175% that of triplePrep and 500% that of AllPrep. triplePrep, therefore, outperforms AllPrep in DNA and protein extraction, but the reverse is true for RNA extraction. Neither kit provides performance equal to that of optimal methods for all three kinds of macromolecules.

Falconer E, Hills M, Naumann U, Poon S S S, Chavez E A, Sanders A D, Zhao Y, Hirst M, Lansdorp P M. DNA template strand sequencing of single-cells maps genomic rearrangements at high resolution. *Nature Methods* 9;2012:1107–1112.

Sister chromatid exchange occurs when double-strand breaks that arise during replication are repaired by homologous recombination pathways. Although the process is mostly error-free, it can lead to copy-number variation, loss of heterozygosity, and aneuploidy. An elevated incidence of sister-chromatid exchange is an indicator of genome stress and instability. Identifying sister-chromatid exchanges in the daughters produced by division of single cells is therefore of importance but has not hitherto been possible with high-resolution sequencing techniques. The present paper presents a method called Strand-seq, which provides high-resolution maps of sister chromatid exchanges and identifies other indicators of genome instability, such as aneuploidy and copy-number variants. It identifies the original parental DNA template strands in daughter cells by incorporating BrdU into the nascent stand during DNA replication and then selectively degrading the nascent strand by photolysis so that the template strand can be amplified for construction of directional sequencing libraries. The technique is anticipated to be useful for analyzing tumor evolution. It will also assist in orienting unbridged contigs in regions of the genome that are difficult to assemble: the authors report 17 incorrectly oriented segments in the mm9 mouse reference sequence, totaling nearly 1% of the genome.

SMALL MOLECULE ANALYSIS AND METABOLOMICS

Creek D J, Chokkathukalam A, Jankevics A, Burgess K E V, Breitling R, Barrett M P. Stable isotope-assisted metabolomics for network-wide metabolic pathway elucidation. *Analytical Chemistry* 84;2012:8442–8447.

Our knowledge of the complex web of interactions that constitutes intermediary metabolism is surprisingly sketchy, especially in less-intensively studied species. The elucidation of pathways has long relied on stable isotope tracing to determine the fate of individual metabolites. Studies have traditionally been performed by methods targeted toward the detection of anticipated metabolic products, but this approach tends to miss the unexpected. Metabolomics, on the other hand, has refined untargeted

doi: 10.7171/jbt.13-2401-006

methods of analysis. The present paper demonstrates the use of combining the untargeted methods of metabolomics with traditional, stable isotope-tracing methodology. This combination is applied to the metabolism of the procyclic lifecycle stage of *Trypanosoma brucei*. Trypanosomes are cultured using medium, in which 50% of the glucose is uniformly ^{13}C-labeled. Controls are cultured in medium containing unlabeled glucose. Metabolites are analyzed by liquid chromatography (LC)-mass spectrometry (MS)/MS using an Orbitrap mass spectrometer. Under the steady-state culture conditions used, ^{13}C is incorporated into 187 of 588 metabolites putatively identified, including carbohydrates, lipids, nucleotides, amino acids, and other diverse compounds. The labeling patterns allow active pathways to be identified. For example, succinate is predominantly labeled in three of its four carbon atoms, a result consistent with fermentative production from pyruvate via pyruvate carboxylase rather than from acetyl CoA, which would produce 2-carbon labeling. Aspartate and the pyrimidine nucleoside, orotate, are also 3-carbon labeled, indicating derivation from phosphoenolpyruvate via oxaloacetate and subsequent de novo pyrimidine synthesis. The labeling patterns of metabolites derived from acetyl-CoA indicate that glucose is not the primary source of acetyl-CoA. Unanticipated metabolites derived from glucose are identified. They include glycerate and gluconate. Glycerate presumably arises by dephosphorylation of phosphoglycerate, a product of glycolysis, and gluconate probably comes from phosphogluconate, a product of the pentose phosphate pathway, although enzymes catalyzing these reactions have not been identified in *T. brucei*. The methodology is generally applicable for assessing metabolic flux patterns and discovering new pathways.

Dénes J, Szabó E, Robinette S L, Szatmári I, Szönyi L, Kreuder J G, Rauterberg E W, Takáts Z. Metabonomics of newborn screening dried blood spot samples: a novel approach in the screening and diagnostics of inborn errors of metabolism. *Analytical Chemistry* 84;2012: 10113–10120.

Traditional methods for mass spectrometric screening for inborn errors of metabolism are here extended by deployment of nanospray ionization and high-resolution MS to encompass detection of over 400 metabolites in blood. These metabolites include acylcarnitines, amino acids, organic acids, fatty acids, carbohydrates, bile acids, and complex lipids. Metabolites are extracted from dried blood with methanol and subjected to nanospray ionization to reduce suppression effects but without introducing an inconvenient chromatographic separation. Mass spectral analysis is performed in positive and negative ion modes with an Orbitrap mass spectrometer, which is operated routinely at a resolution as high as 100,000. Samples are introduced via chip-based infusion to minimize carryover. Quantification is performed with stable, isotope-labeled internal standards. The untargeted nature of the method allows multivariate pattern recognition to be used for improving the specificity with which inborn errors are identified and for distinguishing them from interferences that may result from medical treatment or damaged samples.

Ellis S R, Ferris C J, Gilmore K J, Mitchell T W, Blanksby S J, in het Panhuis M. Direct lipid profiling of single cells from inkjet printed microarrays. *Analytical Chemistry* 84; 2012:9679–9683.

Mass spectrometric analysis of lipid profiles from single cells is accomplished in this paper by making use of the ability to deposit individual live cells in a microarray format onto a glass substrate using a piezoelectric inkjet print head. Droplets deposited are dried under a stream of nitrogen. The number of cells/droplet follows a Poisson distribution. The number of cells/microarray spot is counted automatically, and spots containing just one cell are marked for mass spectral analysis. Lipids are extracted by depositing an organic extraction solvent onto the selected spots and then aspirating the liquid for chip-based nanoelectrospray ionization. An important feature of the methodology is avoidance of oxidative damage to lipids exposed to air under ambient laboratory conditions. This is achieved mainly by eliminating airflow over the printed cells. Nanoelectrospray is chosen in preference to MALDI, as it provides prolonged times for mass spectral acquisition. With the use of this methodology, lipid profiles characteristic of different cell types are identified.

CARBOHYDRATES AND GLYCOCONJUGATES

Ban L, Pettit N, Li L, Stuparu A D, Cai L, Chen W, Guan W, Han W, Wang P G, Mrksich M. Discovery of glycosyltransferases using carbohydrate arrays and mass spectrometry. *Nature Chemical Biology* 8;2012:769–773.

This paper describes the methodology to streamline and extend the search for new glycosyltransferase specificities among the large number of putative glycosyltransferases whose activities remain unconfirmed and uncharacterized. It combines microarray technology with label-free mass spectrometric assay to establish a pipeline for biochemical characterization of small amounts of enzymes expressed by recombinant techniques. Twenty-four different acceptor sugars are arrayed on a gold substrate coated with tri(ethylene glycol) groups to minimize nonspecific surface interactions. The acceptor sugars are incubated with 85 proteins, each in the presence of seven different activated monosaccharide donors, using four different reaction buffers for each combination, giving a total of

60,000 reactions altogether. The reaction products are then desorbed and ionized for MS by MALDI. Four new glycosyltransferases are identified, one with previously unknown specificity. The methodology is expected to hasten the characterization of this understudied group of enzymes and to provide new enzyme reagents for oligosaccharide synthesis.

Hahne H, Gholami A M, Kuster B. Discovery of *O*-GlcNAc-modified proteins in published large-scale proteome data. *Molecular & Cellular Proteomics* 11;2012: 843–850.

Hahne H, Kuster B. Discovery of *O*-GlcNAc-6-phosphate modified proteins in large-scale phosphoproteomics data. *Molecular & Cellular Proteomics* 11;2012:1063–1069.

The post-translational addition of *N*-acetylglucosamine (*O*-GlcNAc) to serine and threonine residues is increasingly becoming recognized as a dynamic, regulatory modification, affecting signal transduction, transcription, translation, and energy metabolism, among other processes. To identify proteins modified by *O*-GlcNAc, Hahne et al. search published LC-MS/MS datasets acquired by higher collision energy dissociation in hybrid linear ion-trap-Orbitrap mass spectrometers. With the use of a scoring algorithm to detect characteristic *O*-GlcNAc neutral-loss fragment ions, they discover and identify hundreds of peptides not recognized previously as bearing this modification and add nearly 100 proteins to the list of proteins so modified. *O*-GlcNAc modifications appear to be less numerous than phosphorylations, but these two kinds of modification can occur near one another on the same peptide. With the use of a similar retrospective-searching approach, Hahne and Kuster further identify 11 mouse proteins bearing phosphorylated *O*-GlcNAc, a modification previously known only in one protein. As unphosphorylated and phosphorylated *O*-GlcNAc frequently exist on peptides of the same sequence, phosphorylated *O*-GlcNAc is presumed to represent the product of successive modification by an *O*-GlcNAc transferase, followed by a yet-to-be-identified kinase. The functional importance of phospho-*O*-GlcNAc remains to be elucidated.

MASS SPECTROMETRY

Gorka J, Bahr U, Karas M. Graphite supported preparation (GSP) of α-cyano-4-hydroxycinnamic acid (CHCA) for matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) for peptides and proteins. *Journal of The American Society for Mass Spectrometry* 23;2012:1949–1954.

The authors show that when the MALDI matrix compound α-cyano-4-hydroxycinnamic acid (CHCA) is applied to a target plate using the dried droplet method, signal strength and reproducibility are increased substantially if the target plate is precoated with graphite particles. A commercially available graphite powder with particle size 1–2 μm is wiped over the metal target with a Q-tip to form an invisibly thin layer. This supports formation of a thin, more uniform layer of small matrix crystals, probably because the graphite particles act as nucleation centers for matrix crystal growth. The result is an increase in MALDI signal strength of up to 8× for small peptides. The limit of detection of BSA improves 5×. Peptide signal strength also becomes more reproducible.

Roth M, Kim J, Maresh E, Plymire D, Corbett J, Zhang J, Patrie S. Thin-layer matrix sublimation with vapor-sorption induced co-crystallization for sensitive and reproducible SAMDI-TOF MS analysis of protein biosensors. *Journal of The American Society for Mass Spectrometry* 23; 2012:1661–1669.

This paper presents details of how to optimize MALDI mass spectral sensitivity using vacuum sublimation to deposit the matrix, CHCA. The sublimation process is controlled by limiting the amount of matrix solution in the reservoir from which the matrix is transferred so as to ensure that the matrix layer deposited on the target plate is thin. However, the matrix is still unable to support desorption/ionization. Its ability to do so is activated by briefly exposing the coated target to an organic solvent vapor to dissolve the matrix in a controlled manner and then allowing it to recrystallize rapidly upon drying. The form of the matrix crystals can be modified by selecting which solvent to use. This procedure results in >10× improvement in sensitivity compared with deposition of matrix by the dried droplet method. It also improves reproducibility of signal strength.

PROTEINS—PURIFICATION AND CHARACTERIZATION

Colella A D, Chegenii N, Tea M N, Gibbins I L, Williams K A, Chataway T K. Comparison of stain-free gels with traditional immunoblot loading control methodology. *Analytical Biochemistry* 430;2012:108–110.

Gürtler A, Kunz N, Gomolka M, Hornhardt S, Friedl A A, McDonald K, Kohn J E, Posch A. Stain-free technology as a normalization tool in Western blot analysis. *Analytical Biochemistry* 433;2013:105–111.

These two papers present data on the best way to normalize signal strength in Western blotting experiments. They indicate that neither of the two routinely used proce-

dures—counterstaining the gels with a protein dye to estimate total protein loading and Western blotting with an antibody against a constitutively expressed protein—is optimal. A better option is estimating total protein using stain-free gels, which contain trihalo compounds that react with tryptophan residues on proteins when exposed to UV light to produce a fluorescent signal. This signal can be recorded using a charge-coupled device imager, allowing the amount of total protein loaded, the efficiency of transfer to a membrane, and the quality of electrophoretic separation to be assessed. The ability to make these observations allows the experimenter to decide rapidly whether to proceed to Western blotting with the expenditure of time and money this entails.

Grupi A, Minton A P. Capillary Viscometer for fully automated measurement of the concentration and shear dependence of the viscosity of macromolecular solutions. *Analytical Chemistry* 84;2012:10732–10736.

Measuring how concentrated protein solutions flow has practical importance for investigating the non-ideal behavior of macromolecules and for the formulation and delivery of pharmaceuticals. For example, estimating the force required to inject a pharmaceutical preparation through a hypodermic needle of any particular gauge and length is a commonly encountered problem. This paper contains a description of a novel viscometer/rheometer that automatically varies solute concentration and shear rate, enabling viscosity and shear forces to be measured over a large range of concentrations. The instrument produces a concentration gradient using a single syringe pump and a six-way valve. It requires as little as 0.75 mL of the original concentrated solution. Commercial instruments need much more sample or else they do not support automated dilution. Viscosities as high as 500 cP and as low as 1 cP are measured with the new instrument at shear rates between 10 and 2×10^3 s^{-1}.

Shahid S A, Bardiaux B, Franks W T, Krabben L, Habeck M, van Rossum B-J, Linke D. Membrane-protein structure determination by solid-state NMR spectroscopy of microcrystals. *Nature Methods* 9;2012:1212–1217.

Shahid et al. report the use of solid-state NMR spectroscopy techniques to solve the three-dimensional structure of a bacterial integral membrane protein, the 105-residue YadA protein from *Yersinia enterocolitica*. Detergents that are used to isolate membrane proteins frequently interfere with crystallization for X-ray diffraction studies and with the rapid tumbling required for solution-state NMR. The present work overcomes these problems by using the solid-start NMR technique of magic-angle spinning, in which the sample is spun at an angle to the direction of the magnetic field—the magic angle—to sharpen the resonance signals. The protein used for the study was microcrystalline. Its quality was unsuitable for X-ray crystallography, as frequently occurs with membrane proteins. The protein was uniformly ^{13}C- and ^{15}N-labeled for NMR, but the work did not require more complex labeling techniques. The results indicate that solid-state NMR is an increasingly viable method for structural studies of otherwise intractable membrane proteins.

PROTEOMICS

Glatter T, Ludwig C, Ahrńe E, Aebersold R, Heck A J R, Schmidt A. Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *Journal of Proteome Research* 11;2012:5145–5156.

This paper documents the improvement in proteolytic efficiency attained by serially digesting proteomic samples with Lys-C and trypsin rather than with trypsin alone. Lys-C digestion is performed in 6 M urea, then the digest is diluted to 1.6 M urea, and finally, it is digested with trypsin. The results of this procedure are compared with trypsin digestion in the presence of RapiGest detergent and trypsin digestion over an extend period of time. The performance of the various digestion protocols was compared in label-free quantitative studies of a yeast cell lysate. Tandem digestion with Lys-C and trypsin is found to be the method of choice.

Frei A P, Jeon O-Y, Kilcher S, Moest H, Henning L M, Jost C, Pluckthun A, Mercer J, Aebersold R, Carreira E M, Wollscheid B. Direct identification of ligand-receptor interactions on living cells and tissues. *Nature Biotechnology* 30;2012:997–1001.

Frei et al. describe a new methodology for identifying the specific cell-surface receptors to which protein ligands, such as peptide hormones, pathogens, and pharmaceutical proteins, bind. The methodology uses a specially designed trifunctional reagent. The first functional group on the reagent is an *N*-hydroxysuccinimide ester, which is used to conjugate to amino groups on the surface of the ligand molecule. The second functional group is a trifluoroacetyl-protected hydrazine, which reacts with aldehyde groups introduced for tagging purposes into glycan moieties of cell-surface glycoproteins under mild conditions. These two functional groups on the trifunctional reagent cross-link the ligand to its receptor. The protecting group on the hydrazine does not need to be removed to react with aldehydes on cell-surface glycoproteins but prevents the hydrazine from reacting with the succinimide ester on the reagent. Spacers in the reagent are designed for flexibility

and solubility, and the functional groups are chosen to allow the ligand-coupling and receptor-capture reactions to proceed under mild conditions. The third functional group on the reagent is biotin, which is used to capture ligand-receptor conjugates for subsequent identification by MS. The application of this reagent is validated by identifying the receptors for insulin, transferrin, apelin, epidermal growth factor, the therapeutic antibody, Herceptin, and two designed ankyrin repeat proteins. It is then used to identify five cell-surface proteins that mediate the binding of the vaccinia virus to cells. The methodology is hoped to provide a general and convenient method for identifying cell-surface receptors.

Eichelbaum K, Winter M, Diaz M B, Herzig S, Krijgsveld J. Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nature Biotechnology* 30;2012:984–990.

The global study of proteins secreted by cultured cells is generally approached in one of two ways, neither of which is satisfactory. Serum-containing growth medium may be used for cell culturing, but this necessitates separating the highly abundant serum proteins from sometimes very small amounts of the secreted proteins of interest. Alternatively, serum-free culture medium may be used, but this approach risks perturbing the assemblage of proteins that the cells secrete. The present study overcomes these problems by using click chemistry to capture newly synthesized proteins and using pulsed stable isotope labeling with amino acids in cell culture (pSILAC) to quantify the secreted proteins. Capture of secreted proteins is accomplished by pulse-labeling newly synthesized proteins with azidohomoalanine, an azide-bearing analog of methionine. The azide group is reacted with an alkyne-functionalized agarose resin by 1,3-cycloaddition (click chemistry) to couple newly secreted proteins to beads. In pSILAC, parallel populations of cells are cultured for a limited time with differentially isotope-labeled lysine and arginine ("intermediate" or "heavy" mass) to allow newly synthesized proteins to be distinguished from pre-existing ones by MS and to quantify the proteins that are synthesized. The methodology is used to compare the proteins secreted by various primary cells and cell lines derived from them. The data show extensive differences in the proteins secreted by primary cells and cell lines, indicating that caution should be exercised in extrapolating results of studies with cell lines to behavior of normal cells in vivo. The present methodology is sensitive enough to permit primary cells to be studied directly. Also studied is the kinetics of protein secretion by macrophages following stimulation by LPS and the effects of serum starvation on protein secretion.

IMMUNOCHEMISTRY AND VACCINES

Petsch B, Schnee M, Vogel A B, Lange E, Hoffmann B, Voss D, Schlake T, Thess A, Kallen K-J, Stitz L, Kramps T. Protective efficacy of in vitro synthesized, specific mRNA vaccines against influenza A virus infection. *Nature Biotechnology* 30;2012:1210–1216.

New seasonal influenza vaccines are designed annually based on predictions of which strains are likely to come into circulation. However, current methods for the production of influenza vaccines, which involve producing antigenic material in eggs or cultured cells, are costly, slow, and variable in yield. If a strain comes into circulation unexpectedly, mounting an effective response against it is nearly impossible. Petsch et al. demonstrate for the first time the use of mRNA vaccines, which may provide a way to resolve these problems. They synthesize mRNAs for intradermal immunization of mice, ferrets, and pigs against influenza A H1N1, H3N2, and H5N1 viruses and demonstrate immunogenicity and/or protection against viral infection. Importantly for preventative application in humans, protective immunity to infection is observed even in very young and very old animals. The vaccines elicit B cell- and T cell-mediated immunity and work against the hemagglutinin, neuraminidase, and nucleoprotein (NP) components of the virus. Responses against the NP are particularly interesting, as this antigen is conserved across virus strains and therefore, might provide a means to elicit cross-protective responses. Nucleic acid vaccines are amenable to rapid production, simplified quality control, and ready scale-up for making millions of doses. RNA is preferable to DNA, as RNA is short-lived in vivo, thus reducing the likelihood of genome integration. RNA is also recognized by innate pattern-recognition receptors, thus reducing the need for adjuvants. Previous attempts by others to make functional RNA vaccines have failed. Petsch et al. attribute their success to protection of the RNA against nuclease digestion by complexation with protamine and to optimizing guanine-cytosine content and inclusion of untranslated region sequences. It is hoped that RNA vaccines may prove effective for protection against infection in human populations.

FUNCTIONAL GENOMICS AND PROTEOMICS

Lin C Y, Lovén J, Rahl P B, Paranal R M, Burge C B, Bradner J E, Lee T I, Young R A. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* 151;2012:56–67.

Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, Wang R, Green Douglas R, Tessarollo L, Casellas R, Zhao K, Levens D. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* 151;2012: 68–79.

Lovén J, Orlando D A, Sigova A A, Lin C Y, Rahl P B, Burge C B, Levens D L, Lee T I, Young R A. Revisiting global gene expression analysis. *Cell* 151;2012:476–482.

MYC is a classic example of an oncogene: the expression of its protein product, c-Myc, is up-regulated by growth-factor stimulation. Mutational changes resulting in c-Myc up-regulation independently of growth-factor stimulation are associated with uncontrolled, cancerous cell growth. The c-Myc oncoprotein is a transcription factor. Its oncogenic activity has long been ascribed to an ability to stimulate the transcription of specific target genes involved in cell growth. Lin et al. and Nie et al. re-evaluate how c-Myc works. Independently, the two groups show that when c-Myc expression increases, it does not up-regulate specific target genes but rather, increases the expression level of the majority of genes that are already transcriptionally active. It does this by increased binding to core promoters, not by binding to previously unoccupied, c-Myc-specific enhancers. It therefore amplifies existing gene expression programs but does not initiate cell proliferation by turning on specific target genes in the manner of a master-regulator. This new finding explains the broad range of effects attributed to c-Myc and the quantitative variation in the effects of c-Myc stimulation from one tumor cell type to another. The appearance of specificity is probably largely an artifact of normalizing measurements of change in gene expression to total RNA levels, which are themselves undergoing global increase. In a primer on methods of gene expression analysis, Lovén et al. draw attention to the possibility that similar misinterpretation has occurred in other gene expression studies and might even be widespread. They advocate normalizing to spike-in mRNA controls to avoid this problem in the future. Meanwhile, there is widespread concern about the integrity of global gene expression studies completed to date.

Lindow M, Vornlocher H-P, Riley D, Kornbrust D J, Burchard J, Whiteley L O, Kamens J, Thompson J D, Nochur S, Younis H, Bartz S, Parry J, Ferrari N, Henry S P, Levin A A. Assessing unintended hybridization-induced biological effects of oligonucleotides. *Nature Biotechnology* 30; 2012:920–923.

Lindow et al. summarize recommendations made by a working group concerned with the safety of oligonucleotide drugs on how to address issues related to off-target interactions by these pharmaceuticals. They recommend that the sequences of antisense oligonucleotides or small interfering RNAs should be selected using appropriate in silico search strategies to avoid binding to unintended target sequences. The expression of any off-target RNAs in locations where drugs accumulate should be anticipated, and possible functional consequences of such binding to predicted, undesired targets should be monitored specifically. The existence of off-target interactions does not preclude the use of RNA drugs but like small-molecule drugs, may limit the window of concentrations within which undesirable side-effects can be avoided. At the same time, effects mediated by interactions not a result of base-pairing, such as binding to pattern receptors of innate immunity, should also be monitored. In addition to their therapeutic application, these recommendations will be of interest to investigators seeking to maintain specificity when using oligonucleotide reagents in an experimental context. The precautions are relevant to ensure correct interpretation of experimental results.

Slavoff S A, Mitchell A J, Schwaid A G, Cabili M N, Ma J, Levin J Z, Karger A D, Budnik B A, Rinn J L, Saghatelian A. Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nature Chemical Biology* 9;2013:59–64.

Peptides that are encoded by previously uncharacterized, short open reading frames have been described before. The present paper identifies a further 86 of these—an unprecedented number. These new peptides were discovered by LC-MS analysis of a pool of peptides <50 aa in length that were isolated from mammalian K562 cells. This length is below the size threshold normally screened in proteomic studies. Peptides shorter than 8 aa and sequences closely related although not identical to ones in known proteins were excluded. The new peptide sequences were compared with RNA-Seq data from K562 cells to confirm that they are not derived from longer open-reading frames (>450 nucleotides) or from mutated versions of open-reading frames that have already been annotated. The abundance range of these peptides falls within that of proteins, They are shown to arise from "non-coding" RNAs or multicystronic RNAs. Translation is often initiated from non-AUG start codons. The number and abundance of such species indicate a previously undemonstrated level of complexity in the proteome.

Abyzov A, Mariani J, Palejev D, Zhang Y, Haney M S, Tomasini L, Ferrandino A F, Rosenberg Belmaker L A, Szekely A, Wilson M, Kocabas A, Calixto N E, Grigorenko E L, Huttner A, Chawarska K, Weissman S, Urban A E, Gerstein M, Vaccarino F M. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* 492;2012:438–442.

The widespread occurrence of genomic variants that result from somatic mutation among the nonmalignant cells of healthy individuals has long been known from karyotypic studies but is still an under-appreciated phenomenon. This paper highlights the phenomenon in modern terms by studying copy-number variants (CNVs) in

induced pluripotent stem cell (iPSC) lines derived from human primary skin fibroblasts. Twenty iPSC lines are derived from seven individuals and are subjected to high-throughput sequencing. Extensive somatic mosaicism for CNVs is observed. At least one-half of the variants is shown by PCR amplification with diagnostic primers, to be present at low frequency in the parental fibroblasts and to have been brought to light by clonal expansion in the process of creating the cell lines. Thirty percent of primary fibroblasts are estimated to contain large CNVs, indicating the extent of genomic heterogeneity among somatic cells. The authors propose induction of pluripotent stem cells to be one way in which this phenomenon can be investigated further. They also draw attention to the need for caution when interpreting the results of studies in which individual transformed cell lines are used to infer how diseases with complex genetic origins arise.

CELL BIOLOGY AND TISSUE ENGINEERING

Wong A P, Bear C E, Chin S, Pasceri P, Thompson T O, Huan L-J, Ratjen F, Ellis J, Rossant J. Directed differentiation of human pluripotent stem cells into mature airway epithelia expressing functional CFTR protein. *Nature Biotechnology* 30;2012:876–882.

Successful in vitro-directed differentiation of human embryonic stem cells (hESC) into functional airway epithelia is reported. Starting with hESC or iPSC, the authors execute a protocol of timed treatments with selected exogenous growth factors designed to mimic in vivo developmental pathways. They then culture the differentiated cells at an air-liquid interface. This procedure promotes the formation of patches of polarized epithelial cells that are coupled by tight junctions, express lung markers, and display active transport by the cystic fibrosis transmembrane conductance regulator (CFTR) protein. The authors proceed to create airway epithelial-type cells from fibroblasts derived from three subjects with the CFTR mutation, F508Δ. This mutation produces a trafficking defect that is the most common cause of cystic fibrosis. Treatment of the resulting cells with a chemical chaperone resulted in increased trafficking of the mutant protein to the cell surface, although it did not enhance transporter function. The present differentiation protocol generates heterogeneous epithelial cell cultures, but with improvements, the process is expected to yield cells that are useful for future disease modeling and drug testing.